

MULTIDIMENZIONÁLIS ADAPTÍV TESZTELÉS ALKALMAZÁSA NAGYMINTÁS MÉRÉSI PROGRAMOKBAN

Andreas Frey

Leibniz-Institut für die Pädagogik der Naturwissenschaften (IPN)

A nagymintás mérések, mint például a PISA, TIMSS vagy a PIRLS eredményei számos országban az oktatási rendszerek hatékonyságára irányították a figyelmet, valamint mind intenzívebb és eredményesebb egyeztetéseket eredményeztek az oktatás területén. Mindazonáltal, a kutatások értékes eredményeiért meglehetősen borsos árat kell fizetnünk, mivel a tipikusan tanulók ezreiből álló nagymintákat a résztvevő országok mindegyikében órákig tartó tesztekkel kell felmérni. A felmérés költségeinek csökkentése, illetve a tanulók és iskolák hosszú távú együttműködésének biztosítása érdekében szükséges megvizsgálnunk azt, hogy milyen lehetőségek állnak rendelkezésünkre a tesztelési eljárások hatékonyságának növelésére úgy, hogy közben a lehető legpontosabb és könnyen értelmezhető eredményeket kapjuk.

Az egyik lehetőség arra, hogy a mérés hatékonyságát annak pontosságának csökkentése nélkül növeljük a számítógépes adaptív tesztelésben rejlik (CAT).

A CAT a látens képességek értékelésének egyik speciális módja, amelyben a vizsgázó számára kiválasztott teszt itemek a vizsgázó előző teszteken adott válaszain alapulnak. E kiválasztási folyamat célja, hogy a vizsgázó számára személyre szabott itemeket tudjunk biztosítani, így növelve az item válaszokból kinyert információk mennyiségét. A hagyományos tesztekhez viszonyítva, amelyek adott számú itemet tartalmaznak, adott sorrendben (fixed item test, FIT), a CAT használatával az itemek számát átlagosan felére lehet csökkenteni, anélkül, hogy a mérés pontossága sérülne. A mérés hatékonyságából eredő előnyöket tovább növelhetjük, ha a CAT hagyományos egydimenziális megközelítését multidimenzió-
nális adaptív teszteléssé (MAT) bővítjük ki.

A prezentáció egy széleskörű szimulációs vizsgálat eredményit mutatja be, amelynek középpontjában a MAT mérési hatékonyságát hasonlítjuk össze a CAT és FIT által nagymintás mérés során nyújtott lehetőségekkel.

Öt dimenziót használva, amelyek egymással való korrelációja $r=0,85$, a MAT mérési hatékonysága (amelyet a felhasznált itemek számának és a hibaérték négyzetgyökének fordítottan arányos átlaga/ számtani közepe ad) 1,3-szor magasabb volt, mint a CAT esetében, valamint 3,7-szer magasabb volt egy véletlenszerűen item választáshoz képest. Így CAT-hoz és a FIT-hez képest a MAT-tal jelentős mértékben növelhetjük a mérési hatékonyságot. Gyakorlatilag a vizsgázónak adott itemek átlagát 73%-kal lehet csökkenteni anélkül, hogy a mérés pontossága sérülne. Ezért a MAT egy ígéretes eszköznek tűnik ahhoz, hogy jelentősen csökkentsük a nagymintás mérésekkel járó erőfeszítéseket.

Egy szimulációs kutatás eredményein keresztül mutatjuk be a MAT nagymintás tesztekben történő alkalmazásának gyakorlati hasznát.